Mobility Access and Transit Infrastructure Impacts on Socioeconomic Disparity
Applying Comparative Data Analytics to Outputs and Outcomes of Infrastructure Projects

Research Thesis

Presented in partial fulfillment of the requirements for graduation
With research distinction in the undergraduate colleges of The Ohio State University

by

Jaehyun Han

The Ohio State University
December 2020

Project Advisors:
Dr. Rudy L. Hightower, II
Dr. Jos C.N. Raadschelders
John Glenn College of Public Affairs

Abstract

Mobility not only means freedom to move, but also increase in potentials and access to new opportunities. Humanity has migrated in both micro- and macro-scale throughout history for various reasons, but most notably for survival and pursuit of more livable life. As mobilization is an important factor for increased economic activities, advancing local travel capability of individuals has been stressed in the realm of public policy, as well as the necessity to resolve the issues involved. The relationship between public transit and poverty is proven to exist, but how to more effectively implement transportation network and infrastructure given such issue remains as an ongoing process. Meanwhile, the world continues to evolve and new technologies and initiatives emerge, so do problems. The ancient Silk Road and the recent China-led Belt and Road Initiative (BRI) illustrate humanity's never-ending quest for movement and development and consequentially have brought up many questions and concerns; some of them include if such development of transportation infrastructure will benefit populations disproportionately or indiscriminatory and if it will further escalate the socioeconomic disparity amongst the populations. Thus, the BRI-related transportation projects will likely add another dimension to the current issues regarding transportation and poverty if they fail to address the concerns appropriately.

This research paper first focuses on the case of Columbus, Ohio by examining how its current transportation, especially as the city emphasizes the ongoing city-wide initiative called Smart Columbus, is serving its mission in improving the connection between the vulnerable population and more economic activities. Then, the countries that are involved in BRI are studies for a broader scope of comparative analysis. By analyzing the current status of the transit access for low-income populations as well as the predicted outcomes from the initiatives, this research

aims to suggest some potential solutions that could be beneficial in improving the systems, but also in reducing the socioeconomic disparity, ultimately.

The available data regarding COTA system, BRI-funded transit infrastructure projects, and demographic data in the regions of interest are the main resources for the study. The research will apply data analytics techniques to examine mobility infrastructures in the city and international level.

Acknowledgement

Every individual need someone who would look at the hidden potential within themselves, properly guides them to discover it, and pushes go further beyond what they think their limits are—especially in academia. I cannot thank Dr. Rudy Hightower more for walking through every step with me, from selecting the right topic that aligns with my passion and skills, as well as future career, to advising me to study deeper about the topic. I also thank Dr. Jos C.N. Raadschelders for assisting me with his expertise in the subject matter and being an advisor.

Table of Contents

1. Introduction

Mobility has been inseparable in various aspects in the history of humankind affecting from the most fundamental part like survival to economy. One of the humanitarian examples where movement of people concerns with right to life is the replacement of refugees; in the Middle East, numerous efforts were made to find a stable and secure dwelling place for the 70.8 million displaced people not only within the region, but also across the entire globe, reshaping the demographics of the world and introducing new global social issues in the modern world. In addition, the way in which mobility impacts the socioeconomic aspect of the society includes how individuals make decisions on their residency and all the intersectionality in between, as well as public policies that are designed and implemented to maximize the positive impacts of travel of people and commodities. Mobility means more than just having the freedom to travel, but also indicates being able to access a larger pool of opportunities that may allow individuals to achieve, or at least increase the probability to do so, their potentials and have a better quality of life. The human nature that aspires to advance its lifestyle has encouraged people to migrate to places with more and better economic opportunities, in both short and long distance.

The expansion of home-delivery services, which enlarged its market especially with the outbreak of the COVID-19, added another dimension in the relationship between mobility and economy. In addition to direct travel, the mobility for commercial and industrial purposes is penetrated into people's daily life. Such expanded role of mobility was feasible due to the development of transportation infrastructures that are upheld by advanced technologies and policies implemented by various scale of governments and their cooperation. In the modern world that is highly interconnected and accessible, purposes of travel diversified and transportation as the medium of travel became significant. Thus, transport systems will continue to serve as the key to connect places to places and people to people regardless of distance.

The relationship between transportation and economy, more specifically the influence of public transportation in shaping the economy in variety of settings and places, has been studied by many researchers and practitioners in the field thus far. This suggests that development of transit systems has become a significant part in city planning and public affairs, and further, that it should be questioned if the current public transportation and the related policies are not undermining certain populations, especially if they are projected to potentially benefit the most from it. Public equity should satisfy the needs and desires of the majority and yet meet justice for the minority simultaneously (U.S. Constitution). Therefore, government should target to develop transportation infrastructure that is accessible and beneficial to as many individuals as possible and serve the beneficiaries indiscriminately so that economic development can benefit the overall society.

This research paper studies different types of transportation-related innovations in the world that differ in scale and location, but have similarities in nature. For short-distance case, it analyzes the case in Columbus, Ohio, one of the biggest and growing metropolitan cities in the United States. The city of Columbus has started various smart transportation initiatives and still has issues regarding disproportional distribution of equity, as well as economic disparity amongst the neighborhoods. Additionally, the China-led Belt and Road Initiative (BRI) is examined as the international, larger scale of transportation system, by looking at some Asian countries that are contracted with the initiatives. The BRI projects that are the continuum of the ancient Silk Road illustrate humanity’s never-ending impulse for movement and development, while reflecting the world that keeps evolving along new technologies and systems. The BRI has gained attention for its possibility to promote economic integration among the countries across the world (Wang, 2020), however, it consequentially has brought up questions and concerns with respect to its potential downside in benefiting populations disproportionately and further escalating socioeconomic disparity amongst the populations.

Thus, this research examines the performance of transportation systems and how they are related to the economic disparity in the corresponding area. This is done by creating a model that evaluates the economic impact in relation to the demographics of the area, infrastructure types, policies, and such outputs. Additionally, this research examines the effect of the projects and aims to ultimately suggest some potential solutions that could be beneficial in improving the system to reduce socioeconomic disparity. By doing so, this research will serve as an opportunity to learn the importance of policy analysis to reduce failure risks and to implement programs and policies that fulfill the true purpose and mission in serving the community.

In the next chapter, the current conflicting narratives regarding the impact of transportation system on socioeconomic status of communities that are constructed by some of the significant studies are reviewed. Then, in Chapter 3, the methodology that is used to conduct this study is introduced. In Chapter 4, the results and analytics derived from the model created are explained and in Chapter 5, the conclusions and prioritized policy recommendations wrap up this paper by summarizing the significant findings, limitations resulted from this study, and the next steps that should be taken. Then appendix and references are listed in Chapter 6 and 7.

2. Literature Review:

Competing ideas on transportation infrastructure and socioeconomic dis/integration

Comparing to countries such as Singapore and the United Kingdom that took remarkable initiatives and further succeeded in establishing effective transit systems, the United States not only has a larger territorial space and more diverse environments, but also has an exceeding usage of privately-owned vehicles and automobile-friendly infrastructures like intrastate highways and roads. According to the American Road and Transportation Builders Association, as much as 80.3 percent of all personal miles of travel is done by private vehicles while public transit only accounts for 7.4 percent. The current statistics does not make advancement of public transit systems as compelling or even look feasible, but rather, tends to undermine the necessity for public transportation innovation in economic standpoint. However, it is important to note that the heavy usage of automobiles is reinforced by the shortage of alternatives and lack of public spending on transportation infrastructure (Economist, 2011). Additionally, delay of developing public transit will only make it unavailable as an option to potentially alleviate, or at least to contribute in such way, some of the related social issues.

Accessible public transit does more than simply supporting efficient travel of individuals. While nearly 2/3 of not-poor households have 2 or more privately-owned vehicles that enable almost 90 percent of all travels made by the population, almost the 3/4 of poor households in the US have limited access to automobiles, meaning there is only 1 or no vehicle available per household (Giuliano, 2005). Therefore, public transit is an essential basic need for the low-income households and lack of such system, on the other hand, makes the population more vulnerable by having them highly dependent on the hardly available or reliable transport means.

Public transportation not only directly benefits individuals without cars, but also contributes to the entire city, reshaping the demographic distribution by decentralizing poverty. Recent study conducted by Pathak et al. suggests that the inflow of low-income population to the suburbs, which are typically occupied by middle- to upper-class, is feasible with accessible and affordable public transit system (2017). In Atlanta, Georgia, one of the largest metropolitan U.S cities, census tracts with access to public bus transit have a higher proportion of low-income households than tracts without—not only in the city-center area which was already known to have a larger proportion of low-income and public transit availability, but also in suburbs—according to their empirical study. This adds another dimension to the idea of intersectionality of poverty, transportation, and housing that has been used to explain the low-income population's decision-making pattern for residency and how they constrain their lives by putting more weight on accessibility to public transit. It was commonly regarded that the low-income households had to choose affordable housing over living close to work and the consequential opportunity cost for them was increased travel time and dependency on public transportation, even more so if the individuals cannot afford private cars (Ohio Association of Community Action Agencies, 2019). However, this finding suggests that the lower-income population can sparse into the suburbs that have more opportunities and services, given that affordable and accessible public transit systems are established. It also implies public transit system may contribute to alleviating issues regarding gentrifications, job decentralization, etc.

In addition, investing more public spending on public transportation brings stimulus effects and long-term effects, supporting the overall economy directly. The contribution of public transportation sector in economic efficiency and growth includes 22 jobs per million dollars of spending (American Public Transportation Association, Economic Impact of Public Transportation Investment, 2014) whereas the same amount of money spent in other U.S. economic area supports 17 jobs in total on average (American Council for an Energy-Efficient Economy).

On the contrary, Chen and Hayne's multilevel assessment of the relationship between economy and the public transportation infrastructures in the U.S using spatial econometric computable general equilibrium concluded that the regional economic impacts of increase in public transportation infrastructure are positive and yet relatively small (2015). They noted that their result showed relatively smaller impact than previously conducted studies because most investments are now put toward maintenance and improvement, rather than establishment which tends to cause a greater degree of outcome. However, it is not a sufficient explanation as the U.S. public spending for transportation infrastructure maintenance and operation indeed has been declined (and it was already spending far less than European countries) while spending for new construction is more than twice as much per person as Britain (The Economist, 2011). That being said, the study also found that public passenger rail and transit infrastructure have a much larger spillover effect on the regional economic growth at the northeast metropolitan and state level, supporting the significance of having passenger transit infrastructure for the overall economic benefit (Chen and Haynes, 2015). This also implies the significance of demographic characteristics in this type of analysis since the effect is noticeable only in the respective region, however, the study does not explicitly take demographic characteristics into account in their model. This may suggest that there is a part in the economic impact of public transport infrastructure not is not exhaustively explained by their model.

According to the 2019 State of Poverty in Ohio report by Ohio Association of Community Action Agencies (OACAA), Ohio does not have public transportation systems that are affordable or practical for average employees, but especially for the lower-income households. Columbus is ranked at 24 out of 49 U.S metropolitan cities for its accessibility to jobs by various modes of transportation, ranking higher than Indianapolis (36), Kansas City (40), Cincinnati (39) and Cleveland (29) (Accessibility Observatory of University of Minnesota). However, when the actual data and map are looked at, few jobs are accessible via transportation within 30 minutes one-way, making it impractical to rely on as the means to get to work. In the meanwhile, the entire city of Columbus is accessible within 30 minutes by automobiles.

Studies regarding public transportation and economy in Columbus are by far insufficient in explaining the relationship between transportation and economic development in precise statistical language. Even though the common narrative among the studies acknowledges the lack of accessibility for lower-income population, there is no study that analyzes the degree of disparity between automobile and public transportation users.

The economic impact of transportation infrastructure is also applicable in larger geography. The China-led BRI projects established intracontinental transport systems across the Eurasia landmass, along sea and air routes that connect to Africa and the Americas. According to a recent report created by Christoph Nedopil Wang regarding the sustainable transport section of the BRI projects, China's BRI will play a significant role in shaping urban transport motivated by large investments in infrastructure and technological capabilities. As much as 42 percent of the Chinese transport-related investments in the BRI participating countries are going into railways, which is the highest proportion amongst the transport infrastructure sector. Under the Green Travel Action Plan (2019-2022), 12 government ministries and departments agreed to improve urban public transportation systems including subways and bus lanes, as well as general public travel condition. 100 Chinese cities also established transit metropolis plans as a part of the city network and sustainability effort.

Being the largest international project that involves numerous countries and investment from various sources, the BRI has attracted many scholars to study its potential economic impact in the world. One of the common narratives that has been established is that the BRI transport infrastructure projects have a highly significant economic effect. There is a positive economic impact resulted from establishing transport infrastructure that yields to increase in GDP (Yii et al, 2018), that is not limited to China or a region within itself, but also in the neighboring countries that participate in the projects such as Singapore, Thailand, Indonesia, etc. (Li et al, 2017). Such impact is more significant in the case of railway systems. Although Li and other scholars' linear-log regression model does not take spatial spillover effects into account, it implicitly suggests the importance of geographical proximity by the way that the case countries are selected and grouped. However, their conclusion provides insufficient insight about difference of the expected economic development amongst the countries, depending on the current socioeconomic status or geography-specific factors.

Such suspicion is reasonable because Yu et al. suggest that the positive spillover effects caused from the China's transport infrastructure is only applicable at the national level while considerable variance exists at the regional level. Such disparity among the regions is significant because it is contradictory to results made by previously conducted studies in the context. Yu et al. provided their insight on such difference by mentioning that their data include a broader selection of transport infrastructure including railway, which has been highly emphasized by the Chinese central government in recent years and has been known to have significant spillover effects. In essence, the spillover effects vary by subregion when more weights are put into railway infrastructures in the model. They further noted that the economic growth in a region may come at the expense of other, supported by the negative spillover from mobile production factors involved in the analysis.

An interesting finding from a recent study by Wang et al. is that one of the main factors that attract China's overseas projects is lack of labor force in the host country. As infrastructure constructions demand enormous labor input for a long period, such circumstance gives China a perfect opportunity to not only establish a project that brings economic benefit to the country, but also use labor forces from China, instead of using the labor force from the host country. While they mention that there are mutual economic benefits for both China and contracted countries, this is contrary to the common belief about long-term infrastructure projects that is not only about the future economic benefits upon completion of construction, but also about unseen outputs like increase in construction employment, market activities by floating population during the process, etc. Therefore, this finding makes it more necessary to investigate how the BRI projects are impacting the economy disproportionally among the countries involved.

Many other studies have supported the positive spatial spillover effects in global context not in the scope of BRI; Shabani and Safaie's case study performed on the maximum likelihood analysis supports the idea such that development of a rail transport infrastructure in an Iranian province leads to economic growth in neighboring countries. Uzbekistan is proven to have positive spatial spillover effects of railroad infrastructures as well (Yoshino and Abidhadjaev). This suggests the necessity to study whether the BRI projects will align with the current narrative of economic impacts of transportation infrastructures in the global context while there exists an insufficient amount of information regarding the BRI-contracted countries given the conflicting ideas on the economic impact of the public infrastructure projects in different levels of regions.

While many studies paid attention to overall economic impact by focusing on GDP or GDP growth as the measure of economic development, this study focuses primarily on the poverty-related factor due to the nature of topic of interest for this paper; public transit and transportation access are especially more crucial for populations of low-income and poverty. Additionally, many studies acknowledge that the means of transportation have different impacts on the economy; for instance, the scale of economic impacts is more likely to be larger, derived by transport sectors such as water and air transportation that are more often used for commercial or industrial purposes. However, they are not merely considered as a daily means of transportation that is easily accessible to the public, so they are going to set aside for this particular analysis.

3. Methodology

A panel data is a dataset where a panel of subjects are observed at multiple times. The fixed effects panel models are the regression models that remove the bias of unobserved, time-invariant heterogeneity between the subject units by estimating the effects of parameters of interest within an individual over-time (Endsley, 2016, 2019). Such advantage of fixed-effects model implies that a change of a particular predictor is likely a cause of a corresponding change in the response, which becomes useful for causal inference. Including time fixed effects in addition to individual subject-level fixed effects makes the model a two-ways fixed effects model.

Because the datasets for the Columbus and BRI case analyses are collected from different areas over different time period, the basic structure of the two-way fixed effects regression model for panel dataset is adopted:

$$Y_{it} = \alpha_i + X_{it}\beta + \mu_t + \varepsilon_{it} \quad \text{where} \quad \varepsilon_{it} \sim N(0, \sigma_y^2)$$

where *i* represents each individual subject and *t* represents specific time period of observation. β is the weighted average of 1) pooled estimator from the ordinary least squares regression model without any fixed effects, 2) the "within" estimator from individual fixed-effects model (the effect of a change in $X_i$on Y within each individual i); and 3) the "between" effect from time-fixed effects only model. Each X represents the predictor variables as specified in Table 1. μ_t is an intercept term specific to the time period that represents the change over time that affects all across the observation units in the same manner. α_i is the intercept term for each group. ε_{it} is the error term of which the homoscedasticity is assumed to meet the linear regression model assumption (Endsley, 2019).

(1) Columbus, OH

In this model for Columbus MSA, the response variable is the percentage of population below the federal poverty line in individual area *i* for the given year *t* (t = 2009-2018), denoted as $Poverty_{it}$. This model will estimate the population percentage that is living below the federal poverty line given social factors that are projected to influence poverty in a society. This will suggest how the covariates are related to the poverty rate in Columbus MSA. The two-way fixed-effects panel data model fitted for Columbus case study is as follow:

$$Poverty_{it} = \alpha_i + Production_{it}\beta_2 + byPt_{it}\beta_2 + Walking_{it}\beta_3 + Income_{it}\beta_4 + Comauto_{it}\beta_5$$
$$+ \mu_t + \varepsilon_{it} \quad \text{where} \quad \varepsilon_{it} \sim N(0, \sigma^2_{Poverty})$$

Table 1
Columbus, OH - Description of Variables and Sources of Data

Variables	Proxy	Definition	Sources
Poverty	Poverty	Percent of People in Poverty	PolicyMap, Census
Production	Employment rate in production-related field	Estimated percent of employed people age 16 or older in Production, Transportation, and Material Moving occupations	PolicyMap, Census
byPt	Public transportation as the mode for work	Percentage of people who commute to work by public transportation	PolicyMap, Census
Walking	Walking as the mode for work	Percentage of people who commute to work by walking	PolicyMap, Census
Income	Per Capita Income	Pere Capita Income (US$)	PolicyMap, Census
Comauto	The ratio of jobs reachable by public transit to jobs by automobiles	The ratio of average number of jobs within 45-minute transit commute to the average number of jobs within 45-minute automobile travel.	Modified by the author; original data from World Bank

Among the prospective predictors initially, some of them that could have been used for analysis had many missing values which were made up by taking the average of the respective region for imputation. The *Comauto* variable is an aggregated version of the average number of jobs within 45-minute transit commute to the one within 45-minute automobile travel that is created to provide more insightful on the importance of accessibility and transportation for economic opportunities in context. Due to the small size of sample collected, splitting the entire set into train and test set is not a feasible option even though doing so would have been useful evaluating the model performance.

The majority of demographic data used in this research are collected from PolicyMap where the access was gained through the Ohio State University. Much of their data are from the most recent (2014-2018) American Community Survey from the Census Bureau. Geographic-coding data such as FIPS and ZIP code references are from the Bureau of Economic Analysis, U.S. Department of Commerce, as well as the City of Columbus GIS portal.

(2) BRI

The selected countries are the ones used in Yii et al.'s empirical study: Kazakhstan, Kyrgyz Republic, Tajikistan, Turkmenistan, and Uzbekistan from Central Asia; Thailand, Indonesia, Vietnam, and Malaysia from ASEAN; China and Mongolia from East Asia. South Korea was added in order to enrich the diversity of dataset and balance the distribution of countries between higher and lower GDP. Some of the countries—Kazakhstan, Uzbekistan, and Kyrgyz Republic, to name a few—are highly critical for the BRI projects due to the geographical location connecting China and Europe or Russia and have obtained numerous large investments for transport hubs, international corridors, and infrastructures (Lall and Leband, 2019). Especially, Kyrgyz Republic and Uzbekistan have a lower GDP and higher poverty gap than the overall mean of the selected countries. Thus, this empirical study will provide tangible insight on the potential economic impacts of the transit and BRI projects on the countries with pre-existing impoverished economic status.

The two-way fixed-effects regression model using panel data is as follow:

$$Poverty\ Gap_{it} = \alpha_i + X_{it}\beta + \mu_t + \varepsilon_{it} \quad \text{where} \quad \varepsilon_{it} \sim N(0, \sigma^2_{PovertyGap})$$

where the response variable Y_{ij} used in this model, *Poverty Gap*, is the mean shortfall in income or consumption from the poverty line $5.50 a day (counting the nonpoor as having zero shortfall), expressed as a percentage of the poverty line (World Bank), for country *i* at given year *t*. X represents the matrix of the predictor variables as specified in Table 2 and demographic-related predictors such as *GDP*, *Health*, and *HCI* are included as predictors so that the model can tell us the relationship between poverty gap and transportation or BRI-related variables. All other parameters and assumptions are the same as explained in the model specification for Columbus, OH case above.

The model of Yii et al. used a variable for the quantity of goods transported on railways whereas the *Railways* variable in this model is a standardized version for both goods and passengers carried to add the perspective of not only commercial, but also for public-use. Additionally, the model includes a predictor variable called *BRI* which is the log accumulated quantities of BRI contracts and investments for transport sector until the given year *t* in the given country *i* so that the economic impact of BRI is more directly and clearly dealt with in the model.

252 observations from 12 countries are collected with respect to 10 variables as shown in Table 2. Some of the variables contain values that are not available or missing, therefore imputation using the respective mean has been done to be used in modeling. The dataset contains information over 20 years from 2000 to 2020, therefore it is converted into panel data with indices of country and year.

Table 2

BRI -Description of Variables and Sources of Data

Variables	Proxy	Definition	Sources
SI.POV.UMIC .GP	Poverty Gap *(response variable)*	Poverty gap at $5.50 a day (2011 PPP) (%)	World Bank
LP.LPI.INFR. XQ	Logistics Performance Index (Logistics)	Quality of trade and transport-related infrastructure (1: low to 5: high)	World Bank
HD.HCI.OVR L	Human Capital Index (HCI)	Human capital index that quantifies the contribution of health and education to the productivity of the next generation of workers (scale 0-1)	World Bank
NY.GDP.DEF L.KD.ZG	Inflation Rate (Inflation)	Inflation, GDP deflator (annual %)	World Bank
SE.XPD.CSE. ZS	Education Expenditure (Education)	Current secondary education expenditure (% of total expenditure in secondary public institutions)	World Bank
std.railways	Standardized index for railways usage (Railways)	Good transported (million ton-km) and passengers carried (million passenger-km) on railways	Modified by the author; original data from World Bank
log.ny.gdp.mkt p.pp.cd	Log of GDP (GDP)	GDP (PPP, Current international $)	World Bank
log.ie.ppn.tran. cd	Log of Investments (Investments)	Public-private partnerships investment in transport (current US$)	World Bank
log.sh.xpd.che x.pc.cd	Log of Health Expenditure (Health)	Current health expenditure per capita (current US$)	World Bank
log.bri	Log of BRI Contract and Investment Quantities (BRI)	Contract and investment for transport sector (in million $)	American Enterprise Institute

4. Results and Analytics

(1) Columbus, OH

Table 3
Descriptive Statistics for Metropolitan Columbus, Ohio

	Poverty	Production	byPt	Walking	Income	Comauto
Mean	25.24	12.84	4.76	5.03	24,750.00	0.057
Std. Dev.	14.06	6.92	4.14	9.83	11,299.54	0.016
Max	59.30	23.91	17.41	54.52	58,500.00	0.061
Min	0.64	2.19	0.00	0.00	3,393.00	0.000

One of the variables with notable descriptive statistics is *Comauto*. The mean value of *Comauto* is 0.057, meaning that there is about an average of 1 job within 45-minute transit commute for every 17.54 jobs within 45-mintue driving. This indicates that are far more jobs that are reachable by automobiles than by public transit in Columbus MSA. Note that the *Comauto* variable has a very small standard deviation and mean and maximum are very close to each other.

The *Walking* variable has a maximum of 54.52, which means there is an observed instance where more than half of the population studied go to work by foot. It turns out the corresponding area is in Franklin County of ZIP code 43210 and FIPS code 39049, which indeed is the Ohio State University (OSU) main campus area where many students reside. It intuitively makes sense that many college students who work on or near campus would not need automobiles because most of the workplace where they work are nearby. Additionally, it is likely that not a large proportion of them is working full time and making it less of concern, the share of the first-time, full-time undergraduate students at OSU main campus that are Pell Grants recipients is only about 17%, which is very low compared to 34% of all other Ohio institutions (Institute for Higher Education Policy, 2018). This implies that many students at OSU main campus are from mid- to high-income backgrounds, so the respective observation from the college campus area is expected to have low relationship with poverty.

Table 4
Correlation Matrix

	Poverty	Produc-tion	byPt	Walking	Income	Comauto
Poverty	1.00					
Produc-tion	0.36	1.00				
byPt	<u>0.74</u>	0.18	1.00			
Walking	0.43	-0.33	0.28	1.00		
Income	<u>-0.72</u>	<u>-0.56</u>	<u>-0.51</u>	-0.31	1.00	
Comauto	0.38	0.29	0.29	0.13	-0.32	1.00

Table 4 shows that the correlation coefficients of the variables are mostly very close to 0 except *Poverty* and *byPt*, and *Income* and other variables. The correlation coefficient of *Poverty* and *byPt* suggests that the two have a strong, positive relationship. *Income* has a moderately strong negative relationship with *Poverty*, *Production*, and *byPt* as well. The Variance Inflation Factor (VIF) was checked to detect if there is any multicollinearity issue among the predictor variables and none of them has a high value, indicating multicollinearity should not be an issue, as shown in Table 5.

Table 5
Variance Inflation Factor (VIF)

Predictors	VIF
Production	2.5771
byPt	1.4271
Walking	1.9056
Income	2.7982
Comauto	1.2006

Table 6

Result of the Panel-Data Fixed Effects Model

Predictors	Coefficients	Std. Error	t-Statistics	Pr(>|t|)
Production	0.41	0.19	2.19	0.031372 *
byPt	1.62	0.24	6.88	1.55e-09 ***
Walking	0.41	0.12	3.48	0.000831 ***
Income	-0.00	0.00	-2.70	0.008532 **
Comauto	62.96	56.00	1.12	0.26445
R^2		0.7562		
Adjusted R^2		0.7364		
F-statistic		45.69		
RSE		7.293		
Degrees of Freedom		75		

Note: *** indicates statistical significance at 0.00, ** at 0.01, and * at 0.05 level, respectively.

The numerical summary of the panel-data fixed effects model suggests that all the predictors except *Comauto* is statistically significant. The estimated coefficient of *Production* variable indicates that a percent change in the number of people employed in the respective field is related to 0.41% increase in the poverty rate. The two transit access-related predictors, *byPt* and *Walking*, tell us that a percentage increase in the number of people who commute to work via public transportation is related to 2.62% increase in poverty and a percentage increase in the one for walking to 0.50% increase. Thus, the relationship between poverty rate and reliance of non-automobile means is significant as seen by the result. The coefficient of *Income* is precisely -0.0003261 before rounding up. Being very close to zero, it indicates that a change in Income does not have a huge impact on poverty gap. Nonetheless it is still statistically significant, meaning that the relationship is valid, but the degree of change is small based on the estimated coefficient. The statistical insignificance of *Comauto* suggests that there is not sufficient evidence to conclude that there is a relationship between *Comauto* and *Poverty*. A potential cause with such issue is that it was flagged by FIPS code, so it had to be converted and then aggregated to configure into the other dataset with all the predictors that was flagged by ZIP code. This process may have caused some lose in information that could have been meaningful and indeed statistically significant otherwise (see Table 13 in Appendix for the actual values). In addition, note that the standard deviation of *Income* is 0.0001207 (shown as 0.00 after rounding up), which is very small and can indicate a possible issue with the small size of dataset. The R-squared and adjusted R-squared values tell us that about 75.62% of variability in poverty are explained by this model.

(2) BRI

Table 7

Descriptive Statistics for Selected BRI-contracted Asian Countries

	Logistics	Poverty Gap	HCI	Infla-tion	Edu-cation	Railways	Invest-ments	GDP	Health	BRI
Total										
Mean	2.73	18.97	0.58	9.27	53.31	1.00	10.23	26.22	4.75	2.37
Std. Dev.	0.58	19.19	0.19	9.64	45.62	2.98	10.28	2.04	1.14	3.38
Max	3.79	64.90	0.85	59.74	100.00	13.11	20.44	30.79	5.52	9.36
Min	1.86	0.000	0.00	-5.99	0.00	0.000	0.00	22.60	1.79	0.00
Countries with Higher Poverty Gap										
Mean	2.61	33.59	0.60	11.11	44.44	1.83	10.35	26.26	4.06	2.45
Std. Dev.	0.52	16.74	0.06	9.92	45.12	4.05	10.43	2.41	0.84	3.42
Max	3.75	64.90	0.69	47.29	98.17	13.11	23.98	30.79	6.09	9.11
Min	1.96	6.50	0.50	-0.84	0.00	0.00	0.00	22.60	1.79	0.00
Countries with Lower Poverty Gap										
Mean	2.84	4.35	0.56	7.42	62.17	0.17	10.12	26.18	5.43	2.29
Std. Dev.	0.62	5.36	0.26	9.01	44.55	0.21	10.18	1.62	0.99	3.36
Max	3.79	7.15	0.85	59.74	100.00	0.75	21.93	28.43	7.73	9.36
Min	1.86	0.00	0.00	-6.00	0.00	0.12	0.00	22.91	3.28	0.00

Since the parameter of interest is poverty gap of countries, the descriptive statistics for the disaggregated groups separated into higher and lower poverty gap countries are shown in Table 7 in addition to the overall statistical summary of the entire group. Countries with poverty gap higher than the overall mean are classified into the high poverty gap group (China, Indonesia, Kyrgyzstan, Tajikistan, Uzbekistan, and Vietnam) whereas the ones with lower are in low poverty gap group (Kazakhstan, Mongolia, Malaysia, Thailand, Turkmenistan, and South Korea). A notable comparison to make is that the mean GDP for both high and low poverty gap groups are very similar log of GDP and indeed the high poverty gap country group has a higher mean of *GDP* than the one of low gap. Additionally, the *BRI* variable has a higher mean in the higher poverty gap group, meaning that there are more BRI contracts and investments involved in the countries with high poverty gap. This implies that modeling the poverty gap with the *BRI* variable will be able to tell us whether having more weights, represented as millions of dollar worth projects, on the corresponding countries is going to positively impact the pre-existing poverty gap in the host country.

Table 8
Correlation Matrix

	Logis-tics	Poverty Gap	HCI	Infla-tion	Edu-cation	Railways	Invest-ments	GDP	Health	BRI
Logistics	1.00									
Poverty Gap	-0.35	1.00								
HCI	0.45	0.12	1.00							
Infla-tion	-0.49	0.44	-0.18	1.00						
Edu-cation	-0.11	-0.17	0.22	-0.16	1.00					
Railways	0.45	0.00	0.14	-0.18	-0.34	1.00				
Invest-ments	0.53	-0.15	0.25	-0.32	0.44	0.35	1.00			
GDP	0.84	-0.14	0.37	-0.42	-0.06	0.57	0.71	1.00		
Health	0.55	-0.66	0.06	-0.41	-0.18	0.11	0.04	0.46	1.00	
BRI	0.12	-0.05	0.16	-0.20	0.40	-0.20	0.36	0.27	0.17	1.00

Table 9
Variance Inflation Factor (VIF)

Predictors	VIF
Logistics	4.5099
HCI	1.6755
Inflation	1.5332
Education	2.7646
Railways	1.9724
Investments	3.6842
GDP	9.3896
Health	2.0624
BRI	1.4923

Table 8 is the matrix that shows the correlation coefficients of the predictors used in modeling. Although there are some variables with a coefficient higher than 0.50 (underlined in the table) which means that there is a relatively strong relationship between the variables, the degrees of such relationships do not seem extreme. However, since the *GDP* variable seems to have higher correlation coefficients with multiple variables, further check seems necessary. To see if they were significantly problematic, the VIF values are calculated as seen in Table 9. None of the predictors have a VIF value higher than the conventional number 10, it is safe to say that they have no multicollinearity issue and can be used for modeling.

The numerical summary of the model suggests that all the variables involved in the model are statistically significant in explaining the poverty gap in the countries that are involved. The model has -16.78 for the coefficient for *Logistics* variables, meaning a unit increase in the *Logistics* variable is related to a 16.78 unit decrease in poverty gap according to this model. This implies that high quality of trade and transport-related infrastructure is linked to a low poverty gap in the selected countries. *Railways* has a predicted coefficient of -1.28, meaning that for a unit increase in the variable, there is a 1.28% decrease in the poverty gap. The coefficient for *Investments* suggests that an increase of log US dollars toward public-private partnerships investments in transport is related to 0.94% decrease in poverty gap. Surprisingly, *BRI* has a positive coefficient in this model. The increase of log million dollars of BRI contract and investments for transport sector is related to 0.46% increase in poverty gap, fixing for all other variables.

Table 10

Result of the Panel-Data Fixed Effects Model

Predictors	Coefficients	Std. Error	t-Statistics	Pr(>|t|)
Logistics	-16.78	2.30	-7.30	4.24e-12 ***
HCI	19.92	4.04	4.93	1.55e-06 ***
Inflation	0.25	0.07	3.39	0.000813 ***
Education	-0.09	0.02	-3.89	0.000130 ***
Railways	-1.28	0.29	-4.43	1.45e-05 ***
Investments	-0.94	0.14	-6.72	1.27e-10 ***
GDP	10.36	0.86	11.98	< 2e-16 ***
Health	-14.50	0.78	-18.59	< 2e-16 ***
BRI	0.46	0.22	2.09	0.0378 *
R^2			0.7813	
Adjusted R^2			0.7732	
F-statistic			96.05	
RSE			9.141	
Degrees of Freedom			242	

Note: *** indicates statistical significance at 0.00, ** at 0.01, and * at 0.05 level, respectively.

The diagnostic plots are checked for linear model assumptions (See Graph 3 in Appendix). The model has a R-squared and adjusted R-squared value of 0.7813 and 0.7732, respectively, meaning that about 78% of the variability of the poverty gap is explained by the model.

In order to evaluate the model better, 60% of the overall dataset is randomly selected and regarded as a training set to fit the regression model while the rest 40% is used as a test set to evaluate the fit. The result of the model fitted with the training set, as shown in Table 11, suggests a similar manner with the one of model fitted with the entire set; however, the *BRI* variable is now appears to be statistically significant at only 0.1 level. It is possible that it may have caused due to the random sampling.

Table 11
Result of the Panel-Data Fixed Effects Model (Using Training Set)

Predictors	Coefficients	Std. Error	t-Statistics	Pr(>|t|)
Logistics	-18.80	3.08	-6.11	9.05e-09 ***
HCI	24.01	5.42	4.43	1.89e-05 ***
Inflation	0.33	0.10	3.50	0.000632 ***
Education	-0.11	0.03	-3.84	0.000188 ***
Railways	-1.32	0.40	-3.33	0.001108 **
Investments	-0.92	0.19	-4.87	2.95e-06 ***
GDP	10.89	1.12	9.73	< 2e-16 ***
Health	-14.73	1.08	-13.63	< 2e-16 ***
BRI	0.52	0.30	1.77	0.0787 .
R^2			0.7980	
Adjusted R^2			0.7851	
F-statistic			61.89	
RSE			9.344	
Degrees of Freedom			141	

Note: *** indicates statistical significance at 0.00, ** at 0.01, and * at 0.05 level, respectively.

The Mean Error (ME) is -1.5511, Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are 120.90 and 10.9956 respectively, and Mean Absolute Error (MAE) is 8.2481. The values are very small and indicate good fit of the model. This suggests that the model fairly accurately predicts the response and could be useful for future analyses, although the statistical insignificance of the *BRI* remain questionable and should be looked into more carefully.

Because the main interest of this empirical study is on poverty gap, countries with high poverty gap are selected separately to fit the model. When only high poverty gap countries are used in the model, the numerical summary suggests that all the variables except *Education* are statistically significant.

Table 12

Result of the Panel-Data Fixed Effects Model (Countries with High Poverty Gap)

Predictors	Coefficients	Std. Error	t-Statistics	Pr(>|t|)
Logistics	-17.09	5.98	-2.86	0.005085 **
HCI	95.84	15.27	6.28	6.17e-09 ***
Inflation	0.22	0.09	2.45	0.015583 *
Education	0.02	0.02	0.78	0.437140
Railways	-2.89	0.49	-5.89	3.93e-08 ***
Investments	-2.44	0.23	-10.83	< 2e-16 ***
GDP	16.99	1.14	14.90	< 2e-16 ***
Health	-12.59	1.45	-8.66	3.24e-14 ***
BRI	-1.09	0.31	-3.56	0.000541 ***
R^2			0.8164	
Adjusted R^2			0.8021	
F-statistic			57.29	
RSE			7.445	
Degrees of Freedom			116	

Note: *** indicates statistical significance at 0.00, ** at 0.01, and * at 0.05 level, respectively.

The relationships between transport-related variables and poverty gap are more notable in this case. Especially, decrease of 1.09 unit in poverty gap is expected for a unit increase in the *BRI* variable, meaning increase of log million dollar worth of BRI contracts or investments in transport sector is related to 1.09% of decrease in the poverty gap at $5.50 a day for countries with high poverty gap, controlling for all other variables. A unit increase in the Railways variable, which indicates the usage of railways for goods and passengers transport, is related to 2.89% decrease and a unit increase in the log of public-private partnerships investment in transport is related to 2.44% decrease in the poverty gap.

In addition, the positive relationship between *GDP* and poverty gap is notable in this result; for countries that already have a high poverty gap, a unit increase in log of GDP is related to 16.99% increase in the gap, controlling for other variables. This potentially suggests that only looking at the *GDP* or its growth may not be fully capturing the entire economic view in this context.

Noting the improvement compared to the model fitted with the entire dataset, the R-squared and adjusted R-squared values using high poverty gap countries are 0.8164 and 0.8021, respectively. It means about 81.64% of variability in poverty gap is explained by this model using the selected countries.

5. Conclusions and Prioritized Recommendations

Previous studies have proved the positive correlation of China's BRI projects with local economic growth rate and infrastructure development and further, the positive posterior mean of the coefficients for GDP growth rate for low-income countries (Wang et al, 2020). However, the result of this empirical study using panel-data fixed effects model suggests that the BRI projects are also related to increase in poverty gap of the selected countries. This suggests that relationship between economy and BRI-projects can be interpreted very differently depending on the perspectives and where the details and weights are put into between the overall GDP and poverty gap. It further implies that the positive overall GDP growth rate should not overshadow the poverty-related measures in policymaking, especially for sectors like public transit or transit infrastructures that are heavily related to low-income population by nature. That being said, the result also suggests that host countries with pre-existing poverty gap higher than the mean may benefit from having more investments and projects put toward transit by decreasing the gap. Therefore, having a large number of BRI-related investment and contracts in the countries, as they already do, is likely to have a positive impact on both the overall GDP growth and decreasing poverty, benefiting both China for expanding its projects and the host countries. In addition, this adds another dimension to the previously conducted studies which concluded that the speed of local economic development that represented in GDP growth rate of the BRI-contracted countries is significantly positive after the BRI was launched (Wang et al, 2020).

Although it is not plausible to generalize a definite impact for different settings, especially because variables or measurements are used differently in the analyses, the findings at least suggest the significant role of transit infrastructure and access in relation to poverty. As seen in the case of Columbus, the positive relationship between poverty rate and the percentage of people using something other than automobiles (i.e. public transit and walking) is observed. This indicates that the areas with more people using non-automobile modes to go to work has a higher percentage of people in poverty. On the other hand, the negative relationship between poverty gap and transit-related variables such as the quality of transport-related infrastructure, railway usage, and public-private partnership investments in transport implies the significance of having a higher quality or more quantity of such transit infrastructure and system in gauging the poverty gap, as seen in the BRI case.

In addition, this research is significant in the sense that the response variable is poverty unlike previous studies that primarily used GDP and the predictor variables include not only demographic characteristics, but also specific public transit-related and/or BRI-related variables.

The biggest limitation of these analyses is on the small size of available data. The BRI data had a large enough data to make imputations for some of the missing variables, but still there were some variables that did not have any data points for the respective country, which were automatically marked zero. It was even worse for the Columbus case because out of ten counties in the Columbus metropolitan area, only two of them were usable for analysis and even then the dataset was highly imbalanced because there are far more observations for Franklin county than for Delaware County. The discrepancy of units for geographical location amongst the variables made the analysis even more challenging because the attempt to unify the unit and configure the datasets potentially has caused the issue of making some of the variables statistically insignificant that could have been insightful indeed. Additionally, there was no data available for the public infrastructure expenditure or project spending in Columbus MSA in county level, so it was implausible to include such variable in the modeling. Otherwise, the inference of the model could have been richer as seen in the BRI case that could use the BRI project investments and contracts quantity measure as a predictor variable. This supports the importance of data for analysis in public policy setting.

Thus, based on the analytical results obtained from this research, the following are suggested by the author for future policy implementation regarding public transit:

1) More investments and contracts for BRI transit-related projects should be put in the countries with high poverty gap as they are not only beneficial for China by expanding their projects, but also for the host countries by decreasing the poverty gap.
2) The COTA system and bus lines in Columbus should continue to be installed in the areas that are heavily populated with low-income population. This paper is a quantitative research without any qualitative components such as firsthand observations, interviews, or surveys, so it is lacking in explaining the quality of the public transit directly experienced by users, other than describing it figuratively with the quality amount. Thus, improving quality of the system and infrastructure is a tremendous part of public transit experience that is not dealt with in this paper and yet highly crucial.
3) The city of Columbus should improve in making data available and accessible to public for policy analysis. The insufficient amount of data makes it challenging to make a strong statement of recommendations, thus conducting an analysis supported by a larger and better quality of data will be greatly helpful.

For future analysis, it would be useful to integrate spatial lag model that take geographical proxy into account because railways, which are mainly dealt with in the model, are expected to have a spatial spillover effects as previously mentioned. Since the scope of data is already within Columbus and Asia, it is foreseen that the spillover effects would not be extreme or any out of extraordinary, but including proximity element will make the model more precise and accurate.

6. References

Chen, Z. and Haynes, K. (2015). Multilevel assessment of public transportation infrastructure: a spatial econometric computable general equilibrium approach.

Endsley, K. A. (2016). Diagnostics for fixed effects panel models in R. http://karthur.org/2016/fixed-effects-panel-models-in-r.html

Endsley, K. A. (2019). Implementing fixed effects panel models in R. http://karthur.org/2019/implmenting-fixed-effects-panel-models-in-r.html.

Fielding, G. J. (1995). Transit in American Cities. In Geography of Urban Transportation, 2nd ed. (S. Hanson, ed.), Guilford Press, New York, 1995.

Institute for Higher Education Policy (2018). The Ohio State University – Main Campus Equity Snapshot.

Lall, S. and Leband, M. (2019). Who Wins, Who Loses? Understanding the Spatially Differentiated Effects of the Belt and Road Initiative. *Policy Research Working Paper. World Bank Group. Social, Urban, Rural and Resilience Global Practice, April 2019.*

Liu, Y. and Wang, X. (2011). The Empirical Analysis on the Economy Growth of the New Silk Road Transportation Economic Zone in China – A panel data model on the basis of six factors.

Meyer, J. R., and Gomez-Ibanez, J. A. (1981). Autos, Transit, and Cities. Harvard University Press, Cambridge, Mass.

Ohio Development Service Agency, Office of Research (2020). The Ohio Poverty Report, June 2020.

Pathak, R., Wyczalkowski, C., and Huang, X. (2017). Public transit access and the changing spatial distribution of poverty. *Regional Science and Urban Economics.*

The Economist. (2011). America’s transport infrastructure - Life in the slow lane

Yii K., Bee K., Cheam, W., Chong, Y., and Lee, C. (2018). Is Transportation Infrastructure Important to the One Belt One Road (OBOR) Initiative? Empirical Evidence from the Selected Asian Countries.

Yu, N. Jong, M., Storm, S., and Mi, J. (2013). Spatial spillover effects of transport infrastructure: evidence from Chinese regions. *Journal of Transport Geography*

Wang, C. (2020). Belt and Road Initiative & Sustainable Transport. Europe-China Eco-Cities Link.

Wang, M., Liu, J., and Sriboonchitta, S. (2020). A Bayesian Analysis of the Determinants of China’s Overseas Contracted Projects in Countries Along the Belt and Road Initiative.

7. Appendix

Table 13
Columbus, OH – Panel Date

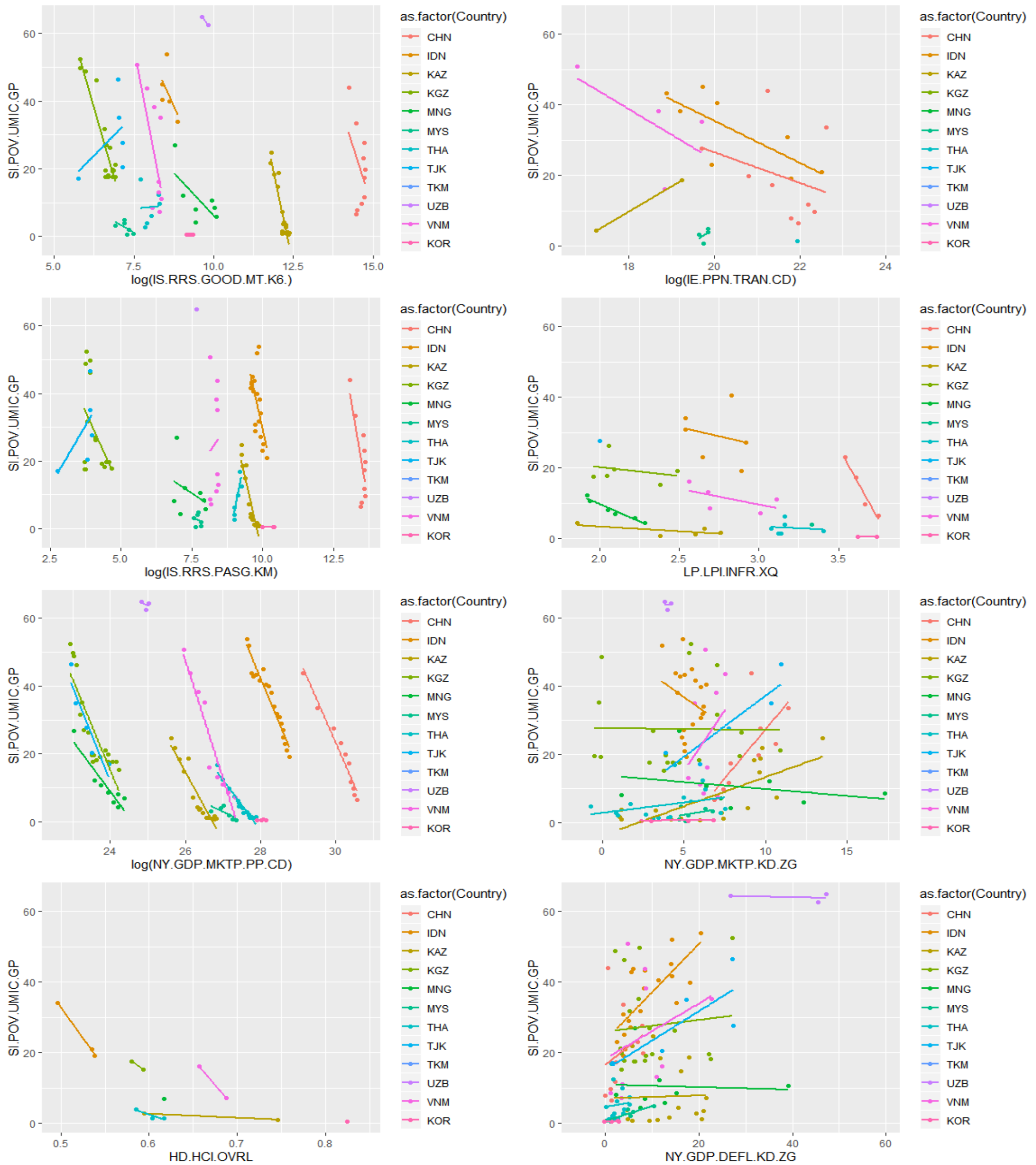
ZIP-FIPS-YEAR	Poverty	Production	byPt	Walking	Income	Comauto
43002-39049-2009	6.82	5.03	0	0	46269	0.060548236
43002-39049-2009	6.82	5.03	0	0	46269	0.060548236
43002-39049-2014	0.64	2.19	0.51	0	45476	0.060548236
43109-39049-2009	2.1	5.05	0	0	22242	0.060548236
43109-39049-2009	2.1	5.05	0	0	22242	0.060548236
43109-39049-2014	6.17	13.89	0	5.71	26921	0.060548236
43201-39049-2009	59.3	6.92	5.81	19.38	15825	0.060548236
43201-39049-2009	59.3	6.92	5.81	19.38	15825	0.060548236
43201-39049-2014	49.89	8.78	5.31	19.84	22601	0.060548236
43202-39049-2009	23.15	5.7	8.5	4.85	26199	0.060548236
43202-39049-2009	23.15	5.7	8.5	4.85	26199	0.060548236
43202-39049-2014	25.15	7.28	7.97	3.7	31386	0.060548236
43203-39049-2009	48.46	12.73	17.41	3.53	12932	0.060548236
43203-39049-2009	48.46	12.73	17.41	3.53	12932	0.060548236
43203-39049-2014	42.83	17.62	16.48	3.2	17711	0.060548236
43204-39049-2009	28.09	19.11	2.97	0.98	20817	0.060548236
43204-39049-2009	28.09	19.11	2.97	0.98	20817	0.060548236
43204-39049-2014	25.4	21.41	1.94	0.56	23746	0.060548236
43205-39049-2009	40.89	14.24	12.75	3.73	17684	0.060548236
43205-39049-2009	40.89	14.24	12.75	3.73	17684	0.060548236
43205-39049-2014	32.64	14.15	8.72	3.56	23865	0.060548236
43206-39049-2009	27.63	8.96	7.49	3.93	30821	0.060548236
43206-39049-2009	27.63	8.96	7.49	3.93	30821	0.060548236
43206-39049-2014	26.07	11.21	6.07	4.25	35313	0.060548236
43207-39049-2009	23.52	23.57	2.57	1.92	18652	0.060548236
43207-39049-2009	23.52	23.57	2.57	1.92	18652	0.060548236
43207-39049-2014	23.39	23.91	3.11	1.12	21048	0.060548236
43209-39049-2009	16.45	5.91	2.22	4.04	36338	0.060548236
43209-39049-2009	16.45	5.91	2.22	4.04	36338	0.060548236
43209-39049-2014	11.62	7.63	2.85	4.13	42820	0.060548236
43210-39049-2009	41.59	2.58	8.89	43.5	3393	0.060548236
43210-39049-2009	41.59	2.58	8.89	43.5	3393	0.060548236
43210-39049-2014	41.84	6.46	10.26	54.52	3976	0.060548236
43211-39049-2009	46.81	18.83	11.72	1.96	11443	0.060548236
43211-39049-2009	46.81	18.83	11.72	1.96	11443	0.060548236

43211-39049-2014	40.39	21.8	10.5	0.75	14555	0.060548236
43213-39049-2009	22.78	19.7	5.23	1.77	22237	0.060548236
43213-39049-2009	22.78	19.7	5.23	1.77	22237	0.060548236
43213-39049-2014	21.59	22.91	5.5	3.02	22493	0.060548236
43214-39049-2009	12.66	5	3.88	1.42	36470	0.060548236
43214-39049-2009	12.66	5	3.88	1.42	36470	0.060548236
43214-39049-2014	10.15	7.99	2.52	1.57	43505	0.060548236
43215-39049-2009	25.9	3.67	3.78	14.8	43646	0.060548236
43215-39049-2009	25.9	3.67	3.78	14.8	43646	0.060548236
43215-39049-2014	19.16	3.74	3.73	15.49	58500	0.060548236
43219-39049-2009	30.89	14.41	3.98	2.42	17684	0.060548236
43219-39049-2009	30.89	14.41	3.98	2.42	17684	0.060548236
43219-39049-2014	33.56	18.35	6.15	1.61	19975	0.060548236
43222-39049-2009	51.51	17.73	5.84	10.85	10764	0.060548236
43222-39049-2009	51.51	17.73	5.84	10.85	10764	0.060548236
43222-39049-2014	50.55	20.35	12.04	1.96	12165	0.060548236
43223-39049-2009	33.14	23.66	2.64	2	14346	0.060548236
43223-39049-2009	33.14	23.66	2.64	2	14346	0.060548236
43223-39049-2014	32.64	23.5	2.86	2.34	16246	0.060548236
43224-39049-2009	27.93	17.43	4.68	1.86	18175	0.060548236
43224-39049-2009	27.93	17.43	4.68	1.86	18175	0.060548236
43224-39049-2014	27.89	23.16	4.53	1.33	19035	0.060548236
43227-39049-2009	26.07	18.26	5.44	1.77	17997	0.060548236
43227-39049-2009	26.07	18.26	5.44	1.77	17997	0.060548236
43227-39049-2014	27.35	23.87	8.88	1.69	19841	0.060548236
43228-39049-2009	26.43	16.09	1.18	1.22	20660	0.060548236
43228-39049-2009	26.43	16.09	1.18	1.22	20660	0.060548236
43228-39049-2014	23.45	19.12	1.03	0.51	23820	0.060548236
43229-39049-2009	23.51	12.25	2.54	1.71	20696	0.060548236
43229-39049-2009	23.51	12.25	2.54	1.71	20696	0.060548236
43229-39049-2014	20.81	18.75	2.27	0.85	21214	0.060548236
43231-39049-2009	17.48	9.27	2.17	0.47	22922	0.060548236
43231-39049-2009	17.48	9.27	2.17	0.47	22922	0.060548236
43231-39049-2014	21.42	14.05	1.41	3.39	23572	0.060548236
43232-39049-2009	23.03	19.41	3.07	2.29	18910	0.060548236
43232-39049-2009	23.03	19.41	3.07	2.29	18910	0.060548236
43232-39049-2014	21.47	23.07	3.41	0.5	20604	0.060548236
43235-39041-2009	7.95	6.07	0.74	1.04	36396	0
43235-39041-2009	7.95	6.07	0.74	1.04	36396	0
43235-39041-2014	7.04	7.06	0.84	0.78	43243	0

43235-39049-2009	7.95	6.07	0.74	1.04	36396	0.060548236
43235-39049-2009	7.95	6.07	0.74	1.04	36396	0.060548236
43235-39049-2014	7.04	7.06	0.84	0.78	43243	0.060548236
43240-39041-2009	5.11	3.61	0	0	34056	0
43240-39041-2009	5.11	3.61	0	0	34056	0
43240-39041-2014	5.57	8.42	0.89	1.66	39971	0

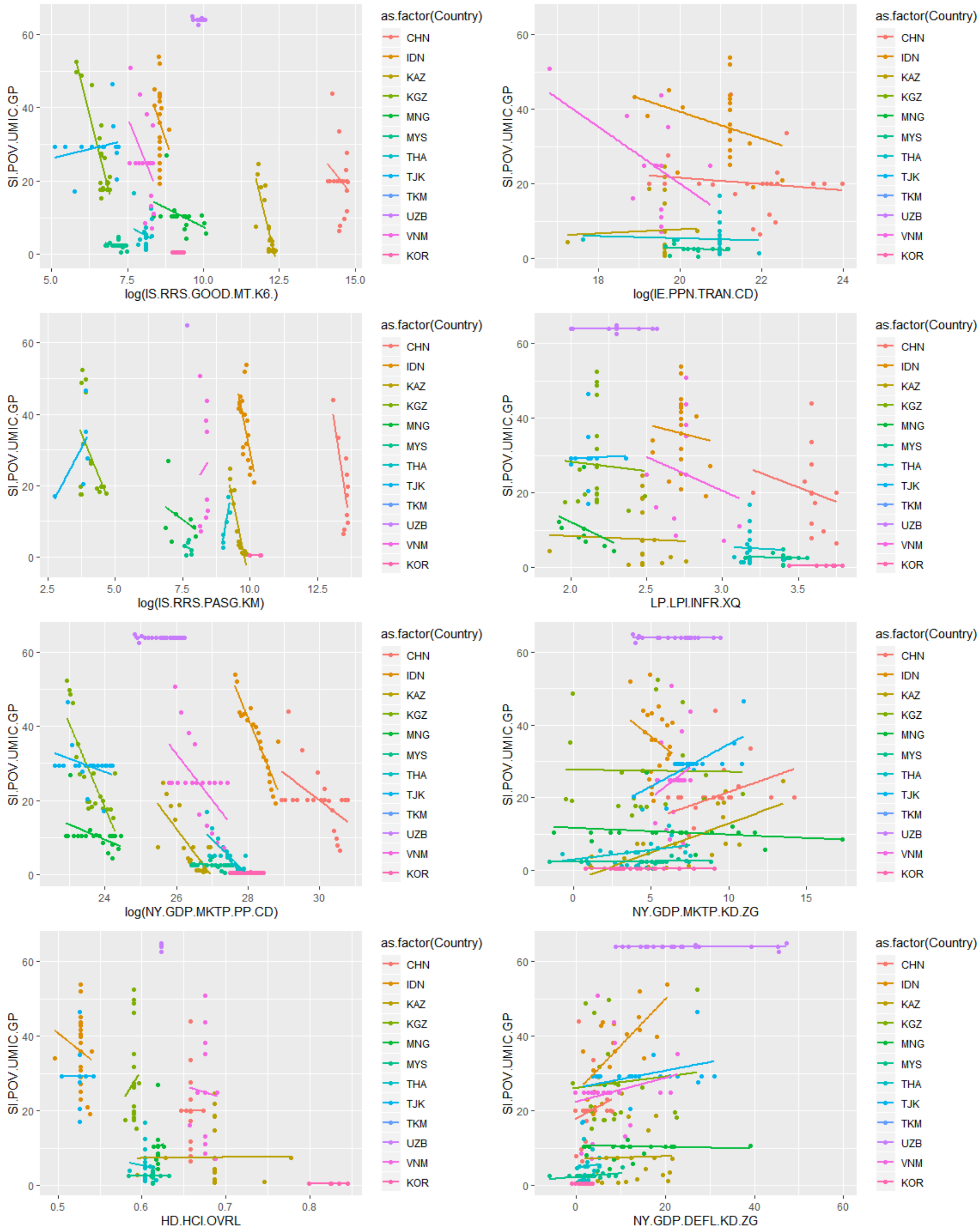
Graph 1

BRI – Exploratory Data Analysis (before imputation)



Graph 2

BRI – Exploratory Data Analysis (after imputation)



Graph 3

BRI - Diagnostic Plots